HYPERSPECTRAL AND MULTISPECTRAL WASSERSTEIN BARYCENTER FOR IMAGE FUSION

Jamila Mifdal, Bartomeu Coll, Nicolas Courty, Jacques Froment, Béatrice Vedel

HYPERSPECTRAL AND MULTISPECTRAL WASSERSTEIN BARYCENTER FOR IMAGE FUSION

Jamila Mifdal[†] [‡] [§]*, Bartomeu Coll* [§]*, Nicolas Courty*[‡]*, Jacques Froment*[†] *and Béatrice Vedel* [†]

[†]Université de Bretagne Sud, LMBA 56000 Vannes, France; [‡]Université de Bretagne Sud, IRISA 56000 Vannes, France
[§]Universitat de Illes Balears Cra. de Valldemossa, km 7.5. Palma de Mallorca, Spain.
jamila.mifdal@univ-ubs.fr

ABSTRACT

The fusion of hyperspectral and multispectral images is a crucial task nowadays for it allows the extraction of relevant information from the fused image. Fusion consists of the combination of the spectral information of the hypespectral image (**h**) and the spatial information of the multispectral image (**m**). The fused image (**f**) has both good spatial and spectral information. In this paper we suggest a new hyperspectral and multispectral image (**h**-**m**) fusion approach based on Optimal Transport (OT) which highlights the idea of energy transfer from the starting images **m** and **h** to the resulting image **f**. The simulations show that the suggested method is effective and compares competitively with other state-of-the-art methods.

Index Terms— Image Fusion, Optimal Transport, Wasserstein Barycenter

1. INTRODUCTION

The fusion of satellite images from different sensors has long been studied and the most famous type of fusion in the literature is the pansharpening. The latter consists of the fusion of a panchromatic image (**pan**) with a multispectral image (**m**), comparative studies about pansharpening methods are available in [14]. Another type of fusion is the hyperspectral pansharpening wich aims at fusing **pan** with a hypespectral image (**h**) [8]. The first method of hyperspectral and multispectral (**h**-**m**) fusion is a wavelet-based technique [7], however, this method depends on the spectral resampling method in case data are missing from **m**, which introduces discrepancies in the reconstructed image.

In this work we present an Optimal-Transport-based hyperspectral and multispectral image fusion. The **h**-**m** fusion has been widely explored throughout methods such as **CNMF** [16] which is based on unsupervised alternate unmixing. In this method, the fused image **f** is found by the combination of **h**'s endmember matrix and the high spatial resolution abundance matrix obtained from **m**. However, **CNMF** method is limited by the fact that each pixel is assumed to be a linear combination of several endmember spectra, and by the fact that both spatial and spectral sensor responses and properties are required. Another approach for **h**-**m** fusion is the **HySure** [12] where the fusion is formulated as a convex optimization problem by using a form of vector Total Variation-based regularization. Similarly to **CNMF**, **HySure** needs both spatial and spectral sensor responses.

Optimal Transport (OT) has been widely used and has applications in many fields such as economy [4], texture synthesis [15], etc. The paper [5] has recently been presented as an interesting tool for data fusion in remote sensing but it does not deal with image fusion. Our paper is the first to present an application of OT in **h**-**m** fusion. After acquisition of two images **m** and **h**, two inverse operators $\widetilde{\mathbf{M}}$ and $\widetilde{\mathbf{H}}$ are applied on **m** and **h** respectively in order to extend these latter to the same domain as **f**. At this stage, classic fusion techniques such as arithmetic and geometric means could be used. However, in our method, which we term **HMWB** for **H**yperspectral and **M**ultispectral **W**asserstein **B**arycenter, we apply the Wasserstein Barycenter (WB) [2] to compute, in a smart and effective way the mean and therefore carry out the fusion task. Empirically, WB has proven to be a powerful tool to compute the barycenter (or the mean) of a set of empirical probability measures. Then we decided to apply it in the case of **h**-**m** fusion where, once again, it proved effective.

The rest of the paper is organized as follows. Section 2 introduces Optimal Transport as a new theory to deal with **h**-**m** fusion. In Section 3 we present the proposed model for **h**-**m** fusion based on an image observation model. Finally, we present the experiments and results of our method (Section 4) followed by a conclusion in section 5.

2. A QUICK OVERVIEW OF OPTIMAL TRANSPORT

Optimal Transport is a mathematical theory which associates a metric between probability distributions. This metric quantifies the main geometric differences between two distributions by measuring the minimal cost of work needed to transport all the mass contained in one distribution onto the other. In the discrete setting, let us consider two discrete probability

measures μ and ν such that (s.t)

$$\mu = \sum_{i=1}^{n} a_i \delta_{x_i} \quad \text{and} \quad \nu = \sum_{j=1}^{m} b_j \delta_{y_j},$$

where $(a_1, \ldots, a_n)$ and $(b_1, \ldots, b_m)$ are probability masses of Dirac located at $X = (x_1, \ldots, x_n)$ and $Y = (y_1, \ldots, y_m)$ respectively. The Wasserstein distance between μ and ν is

$$W(\mu, \nu) = \min_{\phi \in \Pi(\mu,\nu)} < \phi, D_{XY} >, \tag{1}$$

where

$$\Pi(\mu, \nu) = \{\phi \in \mathbb{R}_+^{n \times m} | \phi \mathbb{1}_m = \mu, \phi^T \mathbb{1}_n = \nu\}, \tag{2}$$

$$D_{XY} = [d(x_i, y_j)]_{ij} \in \mathbb{R}^{n \times m}, \tag{3}$$

and d is a distance.

Regularized Wasserstein distance. The numerical implementation of OT algorithms has a high computational cost especially in image processing. For an image of size n, the time complexity is of order $O(n^3 log(n))$ [11]. A way of speeding up the OT computation is the regularization of equation (1) [6].

The regularization has the advantage of making the minimization problem strictly convex, which guaranties the uniqueness of the minimum and the fast computation of the latter throughout the application of Sinkhorn's algorithm. This regularization is carried out by penalizing the entropy of the joint coupling ϕ. The regularized Wasserstein distance is then defined as follows

$$W_\gamma(\mu, \nu) = \min_{\phi \in \Pi(\mu,\nu)} < \phi, D_{XY} > -\gamma E(\phi), \tag{4}$$

where γ is the regularization coefficient and $E(\phi)$ is the entropy defined as

$$E(\phi) = -\sum_{i=1}^{n} \sum_{j=1}^{m} \phi_{i,j}(\log \phi_{i,j} - 1) - \iota_{\mathbb{R}^+}(\phi_{i,j}), \tag{5}$$

where ι is the indicator of $\mathbb{R}_+$ s.t

$$\forall x, \quad \iota_{\mathbb{R}_+}(x) = \begin{cases} 0 & \text{if } x \in \mathbb{R}_+, \\ +\infty & \text{otherwise.} \end{cases} \tag{6}$$

3. THE PROPOSED FUSION MODEL

3.1. Observation model

In the image generation model, the observed **m** and **h** images are respectively supposed to be a spectral and spatial degradation of an ideal image **I**. We suppose that **m** (resp. **h**) has a spatial resolution $n_m \times n_m$ (resp. $n_h \times n_h$) and the number of spectral bands are b_m (resp. b_h). Then **I** has a spatial resolution $n_m \times n_m$ and b_h spectral bands. In fact, we denote that $\mathbf{m} \in \mathrm{R}^{b_m \times n_m^2}$, $\mathbf{h} \in \mathrm{R}^{b_h \times n_h^2}$ and $\mathbf{I} \in \mathrm{R}^{b_h \times n_m^2}$.

Thus, starting from **m** and **h** we would like to have a fused image $\mathbf{I} \in \mathrm{R}^{b_h \times n_m^2}$. The observation model can be written as follows [10]

$$\begin{aligned} \mathbf{h} &= \mathbf{IBS} + \mathbf{N}_\mathrm{h}, \\ \mathbf{m} &= \mathbf{MI} + \mathbf{N}_\mathrm{m}, \end{aligned} \tag{7}$$

where

- $\mathbf{M} \in \mathrm{R}^{b_m \times b_h}$ is the spectral degradation operator which represents the response of the spectral sensor;
- $\mathbf{B} \in \mathrm{R}^{n_m^2 \times n_m^2}$ is a low-pass filter and $\mathbf{S} \in \mathrm{R}^{n_m^2 \times n_h^2}$ is a downsampling operator;
- $\mathbf{N}_h$ and $\mathbf{N}_m$ are **h** and **m** additive white Gaussian noises, respectively. We suppose that both noise are zeros-mean and they are band-dependent. That means, for instance, $\mathbf{N}_h$ can be written as $\mathbf{N}_h$=[$\mathbf{n}_{h,1}$,$\mathbf{n}_{h,2}$,...,$\mathbf{n}_{h,b_h}$] where $\mathbf{n}_{h,i} \sim \mathcal{N}(0, \sigma_{h,i})$.

3.2. Fusion scheme based on Wasserstein Barycenter

As mentioned in the introduction, the fused image **f** is a computed mean between $\widetilde{\mathbf{M}}(\mathbf{m})$ and $\widetilde{\mathbf{H}}(\mathbf{h})$ according to the Wasserstein metric. This leads us to considering the following formulation of the fusion problem

$$\min_{\mathbf{f} \in \Sigma_f} G(\mathbf{f}) = \lambda W_{\gamma_M}(\mathbf{f}, \widetilde{\mathbf{M}}(\mathbf{m})) + (1 - \lambda) W_{\gamma_H}(\mathbf{f}, \widetilde{\mathbf{H}}(\mathbf{h})), \tag{8}$$

f is then the solution of the minimization of the sum of two regularized Wasserstein distances weighted by λ where

- $\Sigma_f = \{\mathbf{f} \in \mathrm{R}_+^{b_h \times n_m^2}, \sum_{i=1}^{b_h n_m^2} \mathbf{f}_i = 1\}$ is the simplex in $\mathrm{R}^{b_h \times n_m^2}$;
- λ is a weight coefficient that favors the spectral or spatial information.

The regularized Wasserstein distance (4) can be recast as a Kullback-Leibler (KL) projection [3]

$$W_\gamma(\mu, \nu) = \min_{\phi \in \Pi(\mu,\nu)} \mathrm{KL}(\phi | \xi), \tag{9}$$

thus the minimization problem (8) can be rewritten as

$$\min\{\lambda \, \mathrm{KL}(\phi_M | \xi_M) + (1 + \lambda) \, \mathrm{KL}(\phi_H | \xi_H), (\phi_M, \phi_H) \in \mathbf{C}\}, \tag{10}$$

where **C** is the intersection of the following four constraints

$$\begin{aligned} \mathbf{C}_{1,M} &= \{\phi_M \in \mathrm{R}_+^{(b_h \times n_m^2)^2}, \phi_M^\intercal \mathbb{1}_{\mathrm{R}^{b_h \times n_m^2}} = \widetilde{\mathbf{M}}(\mathbf{m})\}, \\ \mathbf{C}_{2,M} &= \{\phi_M \in \mathrm{R}_+^{(b_h \times n_m^2)^2}, \phi_M \mathbb{1}_{\mathrm{R}^{b_h \times n_m^2}} = \mathbf{f}\}, \\ \mathbf{C}_{1,H} &= \{\phi_H \in \mathrm{R}_+^{(b_h \times n_m^2)^2}, \phi_H^\intercal \mathbb{1}_{\mathrm{R}^{b_h \times n_m^2}} = \widetilde{\mathbf{H}}(\mathbf{h})\}, \end{aligned}$$

$\mathbf{C}_{2,H} = \{\phi_H \in \mathrm{R}_+^{(b_h \times n_m^2)^2}, \phi_H \mathbb{1}_{\mathrm{R}^{b_h \times n_m^2}} = \mathbf{f}\}$
and (ϕ_M,ϕ_H) is a set of couplings that minimize (10).

Furthermore, for two probability measures p and q, ξ is defined as follows

$$\xi = \underbrace{\exp\left(-\frac{1}{\gamma}\left((x_p - x_q)^2 + (y_p - y_q)^2\right)\right)}_{2D \text{ spatial distances}} \underbrace{\exp\left(-\frac{\alpha}{\gamma}(z_p - z_q)^2\right)}_{1D \text{ spectral distances}},$$

where (x_p, y_p, z_p) and (x_q, y_q, z_q) are two distinct coordinates in a data cube. The parameter α gives the proportion of the spectral Euclidean distance $(z_p - z_q)^2$ with respect to the spatial Euclidean distance $((x_p - x_q)^2 + (y_p - y_q)^2)$.

The minimization of (10) is resolved by applying Bregman projections [3] on the four constraints defined above.

Algorithm 1: HMWB

Input : $\widetilde{\mathbf{M}}(\mathbf{m}), \widetilde{\mathbf{H}}(\mathbf{h}), \boldsymbol{\xi}_m, \boldsymbol{\xi}_h$

Initialization: $\mathbf{v}_M^{(0)} = \mathbf{v}_H^{(0)} = \mathbf{u}_M^{(0)} = \mathbf{u}_H^{(0)} = \mathbb{1}_{\mathrm{R}^{b_h \times n_m^2}}$;

for *i*← *1* **to** *niter* **do**

% Updating the scaling vectors with data (Sinkhorn's iteration)

$\mathbf{v}_M^{(n)} \leftarrow \frac{\widetilde{\mathbf{M}}(\mathbf{m})}{\boldsymbol{\xi}_m * \mathbf{u}_M^{(n)}}, \mathbf{v}_H^{(n)} \leftarrow \frac{\widetilde{\mathbf{H}}(\mathbf{h})}{\boldsymbol{\xi}_h * \mathbf{u}_H^{(n)}}$;

% Computing the barycenter

$\mathbf{f} \leftarrow \exp(\lambda \log(\mathbf{u}_M^{(n)} \odot ((\boldsymbol{\xi}_m * \mathbf{v}_M^{(n)})) + (1 - \lambda) \log(\mathbf{u}_H^{(n)} \odot (\boldsymbol{\xi}_h * \mathbf{v}_H^{(n)})))$;

% Updating the scaling vectors with ***f*** *(Sinkhorn's iteration)*

$\mathbf{u}_M^{(n)} \leftarrow \frac{\mathbf{f}}{(\boldsymbol{\xi}_m * \mathbf{v}_M^{(n)})}, \mathbf{u}_H^{(n)} \leftarrow \frac{\mathbf{f}}{(\boldsymbol{\xi}_h * \mathbf{v}_H^{(n)})}$;

end

Output: **f** (the fused image)

The operation $\odot$ is an elementwise multiplication and "$*$" is a convolution used to speed up the computations [13].

4. EXPERIMENTATION AND RESULTS

4.1. Images generation

In this section we discuss the performance of the **HMWB** algorithm on two different data sets. For each data set, two images **m** and **h** were generated, according to (7), from two reference hyperspectral images which sizes are 128×128×93 taken over the University of Pavia in Italy and over the Japanese city of Chikusei. The images **m** and **h** are both corrupted with Gaussian noises and their sizes are 128×128×4 and 32×32×93 respectively.

4.2. Inverse operators and normalization step

After being generated, **m** is denoised with Wiener filter and **h** is denoised with a Total-Variation based method [1]. In order to apply the fusion algorithm, both images must be on the same domain as **f**. To do this, we introduce two inverse operators $\widetilde{\mathbf{M}}$ and $\widetilde{\mathbf{H}}$ s.t.

$$\widetilde{\mathbf{M}} : \mathrm{R}^{b_m \times n_m^2} \longrightarrow \mathrm{R}^{b_h \times n_m^2} \quad \mathbf{m} \longmapsto \widetilde{\mathbf{M}}(\mathbf{m})$$

$$\widetilde{\mathbf{H}} : \mathrm{R}^{b_h \times n_h^2} \longrightarrow \mathrm{R}^{b_h \times n_m^2} \quad \mathbf{h} \longmapsto \widetilde{\mathbf{H}}(\mathbf{h})$$

where $\widetilde{\mathbf{H}}$ is an interpolation operator and $\widetilde{\mathbf{M}}$ is the solution of the following least square problem

$$\min_{\widetilde{\mathbf{M}} \in \mathrm{R}^{b_h \times b_m}} ||(\mathbf{M} \circ \widetilde{\mathbf{M}})(\mathbf{m}) - \mathbf{m}||^2.$$

Afterwards, the pixel values of $\widetilde{\mathbf{M}}(\mathbf{m})$ and $\widetilde{\mathbf{H}}(\mathbf{h})$ are normalized by the total sum of pixel values of the corresponding image. As a consequence, we can consider both images $\widetilde{\mathbf{M}}(\mathbf{m})$ and $\widetilde{\mathbf{H}}(\mathbf{h})$ as probability measures and apply the **HMWB** algorithm.

4.3. Choice of parameters

The values of the parameters $\gamma_M = 10^{-6}$, $\gamma_H = 10^{-4}$, $\alpha_M = 10^{-3}$, $\alpha_H = 501$ and $\lambda = 0.1$ were fixed for all the experiments. These latter were chosen as the parameters that yielded the best results throughout tests on different data sets. Our method is compared with two state-of-the-art methods: **CNMF** and **HySure** and, with the geometric and the arithmetic mean. In **CNMF** and **HySure** we work with the parameters fixed by their authors.

4.4. Results of the simulations

All experiments are run with MATLAB 2015A. On figures 1 and 2 we notice that, for both data sets, the fused image **f** is visually very close to the reference image **I**. To assess the fusion quality with quantitative results we measure the performance of our algorithm with three quality indices [9]

1. The RMSE is the well known Root Mean Square Error between image **X** and it's estimate $\widehat{\mathbf{X}}$. The closer RMSE to 0 the better the fusion.
2. The Spectral Angle Mapper (SAM) measures the absolute angle between the spectral vector $\boldsymbol{X}_i$ and its estimate $\hat{\boldsymbol{X}}_i$. The closer SAM to 0 the better the fusion.
3. The Cross Correlation (CC) measures the geometric distortion between an image **X** and it's estimate $\widehat{\mathbf{X}}$. The closer CC to 1 the better the fusion performance.

In terms of quality indices, tables 1 and 2 show that for RMSE, SAM and CC our method compares favorably to the other two state of the art methods.

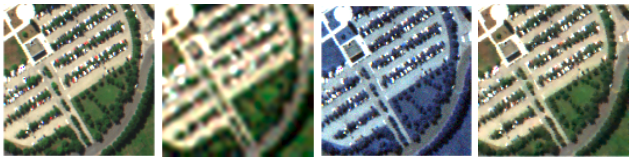

Fig. 1: Pavia data set (false colors). First image (left): ground truth (**I**), second image: **h**, third image: **m**, last image (right): fused image **f**

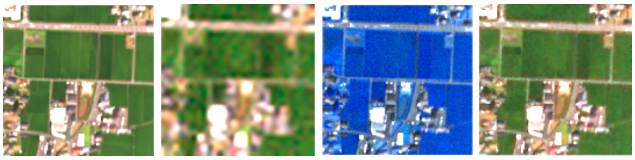

Fig. 2: Chikusei data set (false colors). First image (left): ground truth (**I**), second image: **h**, third image: **m**, last image (right): fused image **f**

Methods	RMSE	SAM	CC
HMWB	$\mathbf{3,1899.10^{-8}}$	**2,0748**	**0,98594**
CNMF [16]	$8{,}6506.10^{-8}$	3,4652	0,89547
HySure [12]	$5{,}6386.10^{-8}$	3,7441	0,95803
G. mean	$1{,}8019.10^{-7}$	14,399	0,94746
A. mean	$1{,}6203.10^{-7}$	12,6038	0,94947

Table 1: Pavia data set. Quality indices of HMWB, CNMF, HySure, Geometric mean and Arithmetic mean

Methods	RMSE	SAM	CC
HMWB	$\mathbf{2,8686.10^{-8}}$	**1,7956**	**0,98694**
CNMF [16]	$3{,}5396.10^{-8}$	1,9274	0,9725
HySure [12]	$4{,}7174.10^{-8}$	2,7484	0,95924
G. mean	$2{,}0564.10^{-7}$	15,9596	0,96315
A. mean	$1{,}8148.10^{-7}$	13,8919	0,96022

Table 2: Chikusei data set. Quality indices of HMWB, CNMF, HySure, Geometric mean and Arithmetic mean

5. CONCLUSION

Throughout this paper, we presented a new approach of hyperspectral and multispectral image fusion based on the application of Optimal Transport. The experimental tests indicate that the results are promising and compares favorably to other state-of-the-art methods.

6. REFERENCES

[1] H. K. Aggarwal and A. Majumdar. Hyperspectral image denoising using spatio-spectral total variation. *IEEE Geoscience and Remote Sensing Letters*, 13(3):442–446, 2016.

[2] M. Agueh and G. Carlier. Barycenters in the wasserstein space. *SIAM Journal on Mathematical Analysis*, 43(2):904–924, 2011.

[3] J.-D. Benamou, G. Carlier, M. Cuturi, L. Nenna, and G. Peyré. Iterative bregman projections for regularized transportation problems. *SIAM Journal on Scientific Computing*, 37(2):A1111–A1138, 2015.

[4] G. Carlier and I. Ekeland. Matching for teams. *Economic theory*, 42(2):397–418, 2010.

[5] N. Courty, R. Flamary, D. Tuia, and T. Corpetti. Optimal transport for data fusion in remote sensing. In *Geoscience and Remote Sensing Symposium (IGARSS), 2016 IEEE International*, pages 3571–3574. IEEE, 2016.

[6] M. Cuturi. Sinkhorn distances: Lightspeed computation of optimal transport. In *Advances in Neural Information Processing Systems*, pages 2292–2300, 2013.

[7] R. B. Gomez, A. Jazaeri, and M. Kafatos. Wavelet-based hyperspectral and multispectral image fusion. In *Aerospace/Defense Sensing, Simulation, and Controls*, pages 36–42. International Society for Optics and Photonics, 2001.

[8] M. M. Khan, J. Chanussot, and L. Alparone. Pansharpening of hyperspectral images using spatial distortion optimization. In *2009 16th IEEE International Conference on Image Processing (ICIP)*, pages 2853–2856. IEEE, 2009.

[9] L. Loncan, L. B. de Almeida, J. M. Bioucas-Dias, X. Briottet, J. Chanussot, N. Dobigeon, S. Fabre, W. Liao, G. A. Licciardi, M. Simoes, et al. Hyperspectral pansharpening: a review. *IEEE Geoscience and remote sensing magazine*, 3(3):27–46, 2015.

[10] R. Molina, A. K. Katsaggelos, and J. Mateos. Bayesian and regularization methods for hyperparameter estimation in image restoration. *IEEE Transactions on Image Processing*, 8(2):231–246, 1999.

[11] O. Pele and M. Werman. Fast and robust earth mover's distances. In *2009 IEEE 12th International Conference on Computer Vision*, pages 460–467. IEEE, 2009.

[12] M. Simões, J. Bioucas-Dias, L. B. Almeida, and J. Chanussot. A convex formulation for hyperspectral image superresolution via subspace-based regularization. *IEEE Transactions on Geoscience and Remote Sensing*, 53(6):3373–3388, 2015.

[13] J. Solomon, F. De Goes, G. Peyré, M. Cuturi, A. Butscher, A. Nguyen, T. Du, and L. Guibas. Convolutional wasserstein distances: Efficient optimal transportation on geometric domains. *ACM Transactions on Graphics (TOG)*, 34(4):66, 2015.

[14] C. Thomas, T. Ranchin, L. Wald, and J. Chanussot. Synthesis of multispectral images to high spatial resolution: A critical review of fusion methods based on remote sensing physics. *IEEE Transactions on Geoscience and Remote Sensing*, 46(5):1301–1312, 2008.

[15] G.-S. Xia, S. Ferradans, G. Peyré, and J.-F. Aujol. Synthesizing and mixing stationary gaussian texture models. *SIAM Journal on Imaging Sciences*, 7(1):476–508, 2014.

[16] N. Yokoya, T. Yairi, and A. Iwasaki. Coupled nonnegative matrix factorization unmixing for hyperspectral and multispectral data fusion. *IEEE Transactions on Geoscience and Remote Sensing*, 50(2):528–537, 2012.